凸优化第 11 周作业

1 预习作业

下节课没有小测。

2 编程题

Problem 1 The *Heavy Ball Method* is a two-step procedure defined by the following state transitions:

$$p^{(k)} = -\nabla f(x^{(k)}) + \beta_k p^{(k-1)} \tag{2.1}$$

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)} \tag{2.2}$$

for some initial points $x^{(0)}$ and $p^{(0)}$, and some positive sequences α_k and β_k. Typically, we just set $p^{(0)} = 0$. This algorithm can be re-written as the iteration

$$x^{(k+1)} = x^{(k)} - \alpha_k \nabla f(x^{(k)}) + \beta_k (x^{(k)} - x^{(k-1)}), \tag{2.3}$$

where the term $x^{(k)} - x^{(k-1)}$ is is referred to as *momentum*. We restrict our attention to the case where α_k and β_k are fixed constants.

We aim to minimize $f : \mathbb{R}^n \to \mathbb{R}$ given by

$$f(x) = \frac{1}{2} x^\top A x - b^\top x + c, \tag{2.4}$$

where A is an $n \times n$ positive definite matrix, b is a vector and c is a constant. We assume that $mI \preceq A \preceq MI$, where $0 < m < M$. This problem has a unique minimizer given by $x^* = A^{-1} b$. We have the chain of equalities

$$\begin{bmatrix} x^{(k+1)} - x^* \\ x^{(k)} - x^* \end{bmatrix} = T(\alpha, \beta) \begin{bmatrix} x^{(k)} - x^* \\ x^{(k-1)} - x^* \end{bmatrix}, \tag{2.5}$$

where $T(\alpha,\beta)=\begin{bmatrix}(1+\beta)I-\alpha A & -\beta I\\ I & 0\end{bmatrix}$.

Consider the following minimization problem:

$$\min f(x_1,x_2)=\frac{1}{2}\left(x_1^2+100x_2^2\right). \tag{2.6}$$

Suppose the starting point is $x^{(0)}=(100,1)^\top$ and we are using the negative gradient as our descent direction.

Consider the **Heavy ball Method** with $\alpha=4/121$ and $\beta=81/121$, then

1. plot the corresponding $x^{(k)}$ on the 2D plane and $f(x^{(k)})$ vs k using semi-log plot.
2. and compare the convergence rate of the Heavy Ball Method to that of standard gradient method by plotting the semi-log plot of $f(x^{(k)})$ vs k.

The algorithms stops when the gradient is less than 10^{-8}.

Problem 2 分别用障碍函数法和原对偶内点法求解下述二次规划问题：

$$\begin{aligned}\min\ & \frac{1}{2}x^\top Px+q^\top x\\ \text{s.t.}\ & Ax=b\\ & x\succeq 0,\end{aligned} \tag{2.7}$$

其中 $x\in\mathbb{R}^n$, $P\in\mathbb{S}_+^n$, $q\in\mathbb{R}^n$, $A\in\mathbb{R}^{m\times n}$, $b\in\mathbb{R}^m$.

1. 障碍函数法要求：
   (a) 阈值误差 $\varepsilon=10^{-8}$.
   (b) 请画出对数对偶间隙 $\log\frac{n}{t}$ 与 Newton 迭代次数 k 之间的关系图.
   (c) 给出原对偶最优解 x^*,λ^*,v^* 和最优值 p^*.

   障碍函数法中参数 μ 建议选取 $\mu=10$ 或者自行选取.

2. 原对偶内点法要求：
   (a) 原误差 $\|r_{\text{pri}}\|_2\le 10^{-8}$，对偶误差 $\|r_{\text{data}}\|_2\le 10^{-8}$，代理对偶间隙 $\hat{\eta}\le 10^{-8}$.
   (b) 分别画出 $\log\hat{\eta}$ 和 $\log\left\{(\|r_{\text{pri}}\|_2^2+\|r_{\text{dual}}\|_2^2)^{\frac{1}{2}}\right\}$ 与 Newton 迭代次数 k 之间的关系图.
   (c) 给出原对偶最优解 x^*,λ^*,v^* 和最优值 p^*.

请**使用“课程作业”附件中提供的数据**求解以上两个问题. 我们给出了 $m=100,n=200$ 时对应的矩阵 P,q,A,b，以及初始点 x_0，λ，v.

3 作业说明

1. 编程作业部分需要撰写报告，包含计算结果/图像/及其分析。报告提交电子版，和代码一起打包提交至网络学堂。提交作业时文件夹中应包含数据文件，保证程序可以直接在文件夹中运行。
2. 编程语言不限，过程需要自己编写，不使用现成的优化器。
3. 第 2 题给出的数据文件可以直接使用 Matlab 打开。
4. 请大家在截止日期前提交作业，过期不候。